Modeling of visual cortical processing to estimate binocular disparity

Christophe Maggia (PhD student, Gipsa-lab, France)
Nathalie Guyader (Assistant Professor, Joseph Fourier university, Gipsa-lab, Grenoble, France),
Anne Guérin-Dugué (Professor, Joseph Fourier university, Gipsa-lab, Grenoble, France),

3D vision is a key process in the spatial understanding of the world. For decades, numerous researches worked on model describing how the brain extracts depth cues from 2D images (monocular and binocular) to create a 3D percept. Among these cues, retinal disparity holds a particular role because it allows stereoscopic vision. According to the objective, vision researchers pursued two different paths. Computer vision models such as matching algorithm can compute reliable depth map, whereas biological inspired models have to describe the mechanism of depth perception from visual input to cortical processes to compute depth map.

The starting point of our work are the physiological and psychophysical studies made on 3D vision, we attempt to build a model of stereoscopic vision. Hence, we used 2D Gabor filters to model the simple and complex cells sensitive to horizontal binocular disparity (Barlow 1967, Daugman 1985). Each of these cells has a preferred disparity and is sensitive to spatial frequency and orientation. It has been shown by Prince et al (2002) that the range of preferred disparities depends on the spatial frequency. We designed a bank of filters in which the distribution of preferred disparity follows the same principle. Moreover, since the stereo-threshold was found to be increasing with the magnitude of disparity inside each spatial frequency channel, the disparity distribution is not uniform. The preferred disparities are closer for small disparity and sparser for larger disparity representing. The magnitude of disparities and spatial frequencies were selected from psychophysical studies (Farrell, 2004) and according to the parameters of an experimental situation. We took the energy model of Ohzawa et al (1986), based on difference of position between the receptive field of simple cells, as a basis since it has been demonstrated that it fits well with the disparity sensitive cells response from V1 in front of most of stimuli. We modified the classical model by normalizing the complex binocular response by the monocular complex response. This step allows the final complex response to be adapted to the spectral content of the stimulus. We took different measures to reduce false matches such as a pooling procedure and an orientation averaging already used by Chen and Qian (2004). As already demonstrated for 2D vision, a coarse-to-fine process seems to be the best way to deal with multiple spatial frequency channels for stereoscopic vision (Smallman 1995, Menz and Freeman 2003). Indeed, the disparity estimation in our model is determined by the most activated cell for each pixel of the image (Winner Takes All, WTA). The first estimation based on low spatial frequencies determines if the estimation will be refined channels depending on its inclusion in the disparity range of the higher spatial frequency channel. If the estimation is included, the coarse to fine process keeps on going whereas it will stay at this stage for estimation out of range. We tested our model both on artificial and natural images. The disparity estimation is quite accurate, but the spatial accuracy is still too low because of the low frequencies. We attempted to enhance it by extracting disparity frontier thanks to a variant of the "2Subunit" model (Haefner and Cumming 2008) and refine depth map by a fine to coarse process that was studied in some studies (Smallman,1995).

This work has been conducted within the Moov3D project, supported by MINALOGIC and IMAGINOVE competitiveness clusters and funded by the Rhône-Alpes regional council.

References

Barlow HB, Blakemore C, Pettigrew JD. The neural mechanism of binocular depth discrimination. *J. Physiol. (Lond.)*. 1967;193(2):327-342.

Daugman JG. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *J Opt Soc Am A*. 1985;2(7):1160-1169.

Prince SJD, Cumming BG, Parker AJ. Range and mechanism of encoding of horizontal disparity in macaque V1. *J. Neurophysiol*. 2002;87(1):209-221.

Farell B, Li S, McKee SP. Coarse scales, fine scales, and their interactions in stereo vision. *J Vis*. 2004;4(6):488-499.

Ohzawa I, Freeman RD. The binocular organization of simple cells in the cat's visual cortex. *J. Neurophysiol*. 1986;56(1):221-242.

Chen Y, Qian N. A coarse-to-fine disparity energy model with both phase-shift and position-shift receptive field mechanisms. *Neural Comput*. 2004;16(8):1545-1577.

Smallman HS. Fine-to-coarse scale disambiguation in stereopsis. *Vision Res*. 1995;35(8):1047-1060.

Menz MD, Freeman RD. Stereoscopic depth processing in the visual cortex: a coarse-to-fine mechanism. *Nat. Neurosci*. 2003;6(1):59-65.

Haefner RM, Cumming BG. Adaptation to natural binocular disparities in primate V1 explained by a generalized energy model. *Neuron*. 2008;57(1):147-158.